-45-

WHAT IS CLAIMED IS:

1.   A system for publishing transcoded media content, comprising:

a publishing service request processor that generates a media provider request based on a requesting client's publishing service request; and

a media provider farm that receives said media provider request from said publishing service request processor and delivers transcoded media content to fulfill said media provider request.

2.   The system of claim 1, wherein said media provider farm comprises:

at least one transcoding server;

a media provider request processor that receives said media provider request and, when indicated by said media provider request, initiates a transcoding task at one of said at least one transcoding server; and

at least one distribution server that distributes said transcoded media content,

wherein said media provider request processor comprises a database containing information about the availability of said at least one transcoding server and said at least one distribution server.

3.   The system of claim 2, wherein said media provider request processor comprises a server load monitor that monitors loading on said at least one transcoding server.

4.   The system of claim 2, wherein said media provider request processor comprises a task load monitor that estimates prospective loading on one of said at least one transcoding server, wherein said prospective loading would be attributable to said transcoding task.

-46-

5.   The system of claim 2, further comprising a media provider request scheduler that generates a batch media provider request that is received by said media provider request processor.

6.   The system of claim 5, wherein said media provider request scheduler comprises a speculative transcoding manager that identifies source media content to be transcoded and identifies a format for the transcoded media content.

7.   The system of claim 2, wherein said media provider farm further comprises a cache that stores transcoded media content for subsequent distribution through said at least one distribution server.

8.   The system of claim 1, wherein said publishing service request processor comprises an intelligent publishing options agent that determines how the transcoded media content is to be published, and generates said media provider request specifying how the transcoded media content is to be published.

9.   A method of providing transcoded media content to a destination client in response to a publishing service request from a requesting client, comprising the steps of:

     (a)   receiving the publishing service request; and

     (b)   when requested transcoded media content is stored in a cache,

          (i)   generating and sending a media provider request that specifies cache access of the requested transcoded media content;

          (ii)   identifying a distribution server;

          (iii)   sending a distribution request to the distribution server; and

          (iv)   delivering the requested transcoded media content from the cache to the destination client via the distribution server.

10. The method of claim 9, further comprising the step of:

(a) when the requested transcoded media content is not stored in the cache in exactly the requested format, but the same content is stored in the cache in a different format usable by the destination client,

(i) generating and sending a media provider request that specifies cache access of the same content in the different format;

(ii) identifying a distribution server;

(iii) sending a distribution request to the distribution server; and

(iv) delivering the same content in the different format from the cache to the destination client via the distribution server.

11. A method of providing transcoded media content to a destination client in response to a publishing service request from a requesting client, comprising the steps of:

(a) receiving the publishing service request; and

(b) given that transcoding capacity is available,

(i) generating and sending a media provider request that specifies on-demand transcoding to satisfy exactly the publishing service request;

(ii) selecting a transcoding server based on server load;

(iii) identifying a distribution server;

(iv) initiating, at the selected transcoding server, a transcoding task corresponding to the publishing service request;

(v) transcoding source media content to produce requested transcoded media content;

(vi) sending the requested transcoded media content to the distribution server;

(vii) sending a distribution request to the distribution server; and

(viii) delivering the requested transcoded media content from the distribution server to the destination client.

12.   The method of claim 11, further comprising the step of:

(ix)   storing the requested transcoded media content in a cache;

performed after step (v).

13.   The method of claim 11, wherein the media provider request of step (i) further specifies prioritized transcoding, and step (b) further comprises the step of:

(ix)   maintaining a list of active transcoding tasks and respective priority information, wherein the priority information is updated by each active transcoding task based on priority criteria.

14.   The method of claim 13, wherein the priority criteria comprises the number of end users that have requested the transcoded media content.

15.   The method of claim 13, wherein the priority criteria comprises information regarding a publishers requirements for publication of the transcoded media content.

16.   The method of claim 11, wherein the media provider request of step (i) further specifies prioritized transcoding, and step (b) further comprises the step of:

(ix)   at the selected transcoding server, killing a transcoding task of lower priority than the transcoding task corresponding to the publishing service request;

performed after step (ii) and before step (v).

17.   The method of claim 11, wherein step (ii) comprises the step of determining server load by measuring current load on the transcoding server,

including any estimations of loads that will be incurred after maturation of immature transcoding tasks.

18.   The method of claim 17, wherein the step of determining server load further comprises the step of receiving an indication of the load of each mature task. .

19.   The method of claim 11, wherein said transcoder server selection of said step (ii) is further based on an estimated load to be created by the transcoding task.

20.   The method of claim 19, further comprising the step of:

(c)   recording historical transcoding task load data at the conclusion of the transcoding task, wherein the historical transcoding task load data comprises source, destination, transcoding server properties, and statistics regarding the load placed on the selected transcoding server by the transcoding task.

21.   The method of claim 20, wherein said step (ii) comprises the steps of:

(A)   determining source, destination, and transcoding server properties;

(B)   searching historical transcode task load data for a closest match;

(C)   calculating a difference between the estimated load of the transcoding task and the closest match; and

(D)   if the difference is sufficiently small, returning a load value of the closest match; otherwise, returning a predetermined conservative load estimate.

22. A method of providing transcoded media content to a destination client in response to a publishing service request from a requesting client, comprising the steps of:

(a) receiving the publishing service request; and

(b) given that transcoding capacity is available,

(i) generating and sending a media provider request that specifies on demand transcoding to produce requested media content in a format different from that requested, but usable by the destination client;

(ii) selecting a transcoding server based on server load;

(iii) identifying a distribution server;

(iv) initiating, at the selected transcoding server, a transcoding task associated with the publishing service request;

(v) transcoding source media content to produce the requested content in the different format;

(vi) sending the transcoded media content to the distribution server;

(vii) sending a distribution request to the distribution server; and

(viii) delivering the transcoded media content from the distribution server to the destination client.

23. The method of claim 22, wherein the media provider request of step (i) further specifies prioritized transcoding, and step (b) further comprises the step of:

(ix) at the selected transcoding server, killing a transcoding task of lower priority than the transcoding task corresponding to the publishing service request;

performed after step (ii) and before step (v).

24. A method of providing transcoded media content to a destination client in response to a publishing service request from a requesting client, comprising the steps of:

(a) receiving the publishing service request; and

(b) directing the requesting client to an origin server that serves the source media content.

25. A method of providing transcoded media content in response to a batch media provider request, comprising the steps of:

(a) receiving the batch media provider request;

(b) selecting a transcoding server based on server load;

(c) sending, to the selected transcoding server, a transcoding task corresponding to the batch media provider request;

(d) transcoding source media content to produce the requested transcoded media content; and

(e) storing the requested transcoded media content in a cache.

26. The method of claim 25, wherein the batch media provider request is generated by a speculative transcoding manager.

27. The method of claim 26, wherein the transcoded media content requested in the batch media provider request is selected on the basis of the popularity of previous transcodes and on their destination types.

28. The method of claim 26, wherein the transcoded media content requested in the batch media provider request is selected on the basis of a publisher's requirement for publication of the transcoded media content.

29. The method of claim 26, wherein said step (d) comprises transcoding in the absence of real time constraints, to produce higher quality transcoded media content than would be produced otherwise. .

30.   The method of claim 25, wherein said step (b) comprises the step of determining server load by measuring current load on the transcoding server, including estimation of any loads that will be incurred after maturation of immature transcoding tasks.

31.   The method of claim 30, wherein the step of determining the server load further comprises the step of receiving an indication of the load of each mature task.

32.   The method of claim 25, wherein said transcoder server selection of said step (b) is further based on an estimated load to be created by the transcoding task.

33.   The method of claim 32, wherein said step (b) comprises the steps of:

    (i)   determining source, destination, and server properties;

    (ii)   searching historical transcode task load data for a closest match;

    (iii)   calculating a difference between the transcoding task and the closest match; and

    (iv)   if the difference is sufficiently small, returning a load value of the closest match; otherwise, returning a predetermined conservative load estimate.

34.   A method for determining how to publish requested transcoded media content, comprising the steps of:

    (a)   receiving a publishing service request from a client;

    (b)   determining an optimum publishing option;

    (c)   generating a media provider request specifying the determined option; and

(d)   sending the media provider request to a media provider request processor.

35.   The method of claim 34, wherein said step (b) comprises the steps of:

(i)   evaluating a set of publishing options, the set comprising:

publication based on matching the requested content with cache contents;

publishing based on near-matching of the requested content with cache contents;

publication based on on-demand transcoding without a priority optimization and without a variance of constraints;

publication based on on-demand transcoding according to a priority optimization and without a variance of constraints;

publication based on on-demand transcoding according to a variance of constraints and without a priority optimization;

publication based on on-demand transcoding according to a variance of constraints and according to a priority optimization; and

publication based on pass-through access to an origin server; and

(ii)   selecting one of the publishing options.

36.   A method of speculatively issuing a batch media provider request, comprising the steps of:

(a)   if transcoding capacity is available, determining whether any pre-encodes have been requested; and

(b)   if any pre-encodes have been requested, issuing a pre-encode batch media provider request, otherwise,

predicting a a future publishing service request, and issuing a corresponding pre-encode batch media provider request.

-54-

37.   A method of ascertaining availability of a server, comprising the steps of:

(a)   receiving, at a database, registration information from the server at fixed intervals, wherein the registration information comprises a list of services that the server is configured to perform; and

(b)   if and only if the servers registration information is received at the most recent interval, concluding that the server is available and performing application level checks to confirm that the server is usable.

38.   A method of shutting down a server in a manner transparent to a client, comprising the steps of:

(a)   issuing a shutdown primitive to the server;

(b)   causing the server to remove its registration information from a database;

(c)   waiting for all currently running tasks to complete; and

(d)   taking down the server.